Ultimate IT Test Preparation Material

NVIDIA
NCA-AIIO
AI Infrastructure and Operations

- **Up to Date products, reliable and verified.**
- **Questions and Answers in PDF Format.**

Full Version Features:

- **90 Days Free Updates**
- **30 Days Money Back Guarantee**
- **Instant Download Once Purchased**
- **24 Hours Live Chat Support**

For More Information:

https://www.testsexpert.com/

- **Product Version**

Latest Version: 6.0

Question: 1

You are evaluating the performance of two AI models on a classification task. Model A has an accuracy of 85%, while Model B has an accuracy of 88%. However, Model A's F1 score is 0.90, and Model B's F1 score is 0.88. Which model would you choose based on the F1 score, and why?

A. Model A - The F1 score is higher, indicating better balance between precision and recall.
B. Model B - The higher accuracy indicates overall better performance.
C. Neither - The choice depends entirely on the specific use case.
D. Model B - The F1 score is lower but accuracy is more reliable.

Answer: A

Question: 2

Which NVIDIA hardware and software combination is best suited for training large-scale deep learning models in a data center environment?

A. NVIDIA Jetson Nano with TensorRT for training.
B. NVIDIA DGX Station with CUDA toolkit for model deployment.
C. NVIDIA A100 Tensor Core GPUs with PyTorch and CUDA for model training.
D. NVIDIA Quadro GPUs with RAPIDS for real-time analytics.

Answer: C

Question: 3

A healthcare company is looking to adopt AI for early diagnosis of diseases through medical imaging. They need to understand why AI has become so effective recently. Which factor should they consider as most impactful in enabling AI to perform complex tasks like image recognition at scale?

A. Advances in GPU technology, enabling faster processing of large datasets required for AI tasks.
B. Development of new programming languages specifically for AI.
C. Increased availability of medical imaging data, allowing for better machine learning model training.
D. Reduction in data storage costs, allowing for more data to be collected and stored.

Answer: A

Question: 4

Which of the following networking features is MOST critical when designing an AI environment to handle large-scale deep learning model training?

A. Enabling network redundancy to prevent single points of failure.
B. Implementing network segmentation to isolate different parts of the AI environment.
C. High network throughput with low latency between compute nodes.
D. Using Wi-Fi for flexibility in connecting compute nodes.

Answer: C

Question: 5

Your AI data center is running multiple high-performance GPU workloads, and you notice that certain servers are being underutilized while others are consistently at full capacity, leading to inefficiencies. Which of the following strategies would be most effective in balancing the workload across your AI data center?

A. Implement NVIDIA GPU Operator with Kubernetes for Automatic Resource Scheduling
B. Use Horizontal Scaling to Add More Servers
C. Manually Reassign Workloads Based on Current Utilization
D. Increase Cooling Capacity in the Data Center

Answer: A

Question: 6

You are tasked with deploying a machine learning model into a production environment for real-time fraud detection in financial transactions. The model needs to continuously learn from new data and adapt to emerging patterns of fraudulent behavior. Which of the following approaches should you implement to ensure the model's accuracy and relevance over time?

A. Continuously retrain the model using a streaming data pipeline
B. Run the model in parallel with rule-based systems to ensure redundancy
C. Deploy the model once and retrain it only when accuracy drops significantly
D. Use a static dataset to retrain the model periodically

Answer: A

Question: 7

An enterprise is deploying a large-scale AI model for real-time image recognition. They face challenges with scalability and need to ensure high availability while minimizing latency. Which combination of NVIDIA technologies would best address these needs?

A. NVIDIA CUDA and NCCL
B. NVIDIA Triton Inference Server and GPUDirect RDMA
C. NVIDIA DeepStream and NGC Container Registry
D. NVIDIA TensorRT and NVLink

Answer: D

Question: 8

A company is using a multi-GPU server for training a deep learning model. The training process is extremely slow, and after investigation, it is found that the GPUs are not being utilized efficiently. The system uses NVLink, and the software stack includes CUDA, cuDNN, and NCCL. Which of the following actions is most likely to improve GPU utilization and overall training performance?

A. Increase the batch size
B. Update the CUDA version to the latest release
C. Disable NVLink and use PCIe for inter-GPU communication
D. Optimize the model's code to use mixed-precision training

Answer: A

Question: 9

In an AI data center, you are responsible for monitoring the performance of a GPU cluster used for largescale model training. Which of the following monitoring strategies would best help you identify and address performance bottlenecks?

A. Monitor only the GPU utilization metrics to ensure that all GPUs are being used at full capacity.
B. Focus on job completion times to ensure that the most critical jobs are being finished on schedule.
C. Track CPU, GPU, and network utilization simultaneously to identify any resource imbalances that could lead to bottlenecks.
D. Use predictive analytics to forecast future GPU utilization, adjusting resources before bottlenecks occur.

Answer: C

Question: 10

You are assisting a senior data scientist in analyzing a large dataset of customer transactions to identify potential fraud. The dataset contains several hundred features, but the senior team member advises you to focus on feature selection before applying any machine learning models. Which approach should you take under their supervision to ensure that only the most relevant features are used?

A. Select features randomly to reduce the number of features while maintaining diversity.
B. Ignore the feature selection step and use all features in the initial model.
C. Use correlation analysis to identify and remove features that are highly correlated with each other.
D. Use Principal Component Analysis (PCA) to reduce the dataset to a single feature.

Answer: C

For More Information – **Visit link below:**

https://www.testsexpert.com/

16$ Discount Coupon: 9M2GK4NW

Features:

- Money Back Guarantee……..……...…. 

- 100% Course Coverage…… 

- 90 Days Free Updates…… 

- Instant Email Delivery after Order……